DOCUMENT RESUME

ED 418 373 CS 013 040

AUTHOR Jones, Lex
TITLE Reading National Assessment.
PUB DATE 1996-03-00
NOTE 11p.; Paper presented at the World Literacy Conference (Philadelphia, PA, March 12-15, 1996).
AVAILABLE FROM Thirty-three selected papers from this conference are available on the "Literacy Online" Website: http://www.literacyonline.org
PUB TYPE Opinion Papers (120) -- Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *British National Curriculum; Foreign Countries; *Literacy; Primary Education; *Standardized Tests; Test Reliability; Test Validity; Testing Programs
IDENTIFIERS Authentic Assessment; England; Wales

ABSTRACT

In England and Wales, a National Curriculum initiated in 1988 was designed to ensure that all schools provided a curriculum which represented different areas of knowledge. The past 20 years has increasingly seen more emphasis on the link between the financial amounts spent on education and subsequent return on this money. The impact of the National Curriculum and the related tests is filtering down the educational ladder. A consequence could well be a preschool National Curriculum. Formal trials of Standard Assessment Tasks (SATs) were conducted on a sample size of over 6,000 7-year-old children from a range of schools across England. However, recognized standards of validity and reliability seem remarkably absent in the model of SATs. What worries the teaching profession is ranking schools on merit from the results obtained across particular subjects on tests which have weaknesses. There seems to be a national isolationism with respect to appropriate approaches to testing literacy. A significant lesson from this experience is the importance of having teachers as part of the process. The revised National Curriculum with "level descriptions" for teaching and learning is a step forward in putting the child at the center of the assessment process. Government recognition that primary teachers need more help with the complexities in assessing children's progress is another welcome step. The wealth of international experience should be harnessed to refine national assessment to alleviate the expensive trial and error initiatives. (Contains 18 references.) (RS)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

LITERACY ONLINE

PROCEEDINGS OF THE 1996 WORLD CONFERENCE ON LITERACY

ED 418 373

# READING NATIONAL ASSESSMENT

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Lex Jones
Middlesex University
Trent Park, Bramley Road
Oakwood, London N14 4YZ

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

L. Jones

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)



| Introduction | The Contemporary Scene | Questions of Procedure | Assessing Changes | National Leagues | Problems | International Understanding | Conclusions | Bibliography |

## INTRODUCTION

The title of the paper may be read in two ways. Such are the nuances of language in allowing for interpretation; it is my intention to achieve both purposes. Language then is a rather creative entity to assess in both its spoken and written forms; some might say an impossible task. But, this has been the situation for many years; standardized tests and informal tests for reading; teacher judgments for writing. At such an important international forum as this World Conference on Literacy it seems to me that sharing approaches for assessing literacy is crucial as approaches to assessment become more contested.

This is particularly needed when the quality of children's reading and writing is increasingly equated with expenditure on education. Perhaps more so when costs of developing and implementing tests are explored and viewed in the context of political ideology. There is also greater awareness of individual rights....and the right for a child to be able to read and write. Taking legal action to challenge schools is an increasing phenomenon so that effective teaching of literacy is evident with recognized standards. While negligence grounds have been the basis of court cases in literacy in the USA for some years, the litigation is gaining momentum in the UK after a 'House of Lords ruling last year upheld the right of two adults and a 16 year old to sue their local education authorities for failing to give them a decent education'. (Scott-Clark & Hymas 1996:7)

It would seem that many countries are currently grappling with the inherent problems in assessment with different formats being tried to suit national, regional and local requirements. One does wonder to what extent this knowledge is shared beyond the confines of the academic journals? Add to this the range of political ideology of developed countries along with the costs of implementing educational policy then a perspective to the complex situations can be appreciated.

While driving to work one morning recently and listening to the radio, my attention was caught by a reading of 'Letters from Over Here' by Raymond Seitz, former US Ambassador. I found the contrasting view between the governing systems of the USA and the UK quite a thoughtful perspective. Power in the UK was seen through the electoral system to be concentrated in those who won; for five years the winning party could exercise its power in pursue its policy into practice with directed purpose and minimal opposition. This was representative of the Westminster parliamentary

CS 013 040



BEST COPY AVAILABLE

system. The USA, on the other hand, had a less concentrated approach to power- almost fragmentary which could control power through restriction. The recent blocking of President Bill Clinton's budget may best illustrate the brakes that can be applied to policy and power. I wonder to what extent such perspectives are reflected through curriculum, assessment and evaluation of both countries. England and Wales in recent years have moved from an autonomous assessment system at primary school level to a highly centralized one, nationally; this does seem to differ from the USA at the federal level.

Identifying literacy standards and gauging these would seem an international problem. In January, 1996 a seminar was held in a city called Reading (yet another meaning of the word) in the county of Berkshire. Sponsored by the British Council and the Government body responsible for curriculum and assessment, SCAA (School Curriculum and Assessment Authority) there were eighteen European countries in attendance discussing issues and concerns coming from national assessment. In Australia, the Commonwealth Government's White Paper, Working Nation (May, 1994) allocated $2.6 million to collect data on literacy by the end of 1996. 'The need for rich, reliable national data on literacy attainment among Australian school students has become a more important issue in recent years.' Australian Literacy Federation(1995:1) There is indeed much we can share internationally to save teachers and children the wasted hours that goes with the experimentation of tests.

In England and Wales, a National Curriculum initiated in 1988 was designed to ensure that all schools provided a curriculum which represented different areas of knowledge. Prior to the introduction of the National Curriculum, schools varied enormously in relation to depth, breadth, focus and expectations. For the first time in many years schools had a single track to follow rather than the highly individual approach which characterized Primary (elementary) education. Secondary education had designated examinations. As these were conducted from selected syllabus determined by various examination boards then a measure of students' learning was evident.

The following paper seeks to examine some of the current issues arising from the mandatory SATs (Standard Assessment Tasks) administered to seven year old children. The SATs are seen by the Chief Executive of the School Curriculum and Assessment Authority in a recent review as being 'firmly on track to deliver full information in future about standards of performance in all our schools'.(Haigh: 1995). What is meant by full information and what determines a standard of performance would seem to be concerns at the heart of the assessment debate. While informal teacher assessment is an essential component, it has not received the development and impetus of mandatory national tests.

## THE CONTEMPORARY SCENE

Expenditure on education remains a contentious issue when equated with quality. The past twenty years has increasingly seen more emphasis on the link between the financial amounts spent on education and subsequent return on this money. As developed societies respond to refine social needs in welfare, health and education, the growing demands on the public purse (coupled with changing work patterns and decreasing numbers in employment) necessitates money going into more public spending baskets. It stands to reason that education is going to have to compete for its share of the financial purse along with other needs in society. Questions of accountability for expenditure are more to the forefront than previously. Chris Woodhead, the HM Chief Inspector (1996:2) in defending the need for external inspections of schools to 'deliver reliable, objective judgments' believes that the 'Government has a perfectly legitimate right to know how the £23 billion of taxpayers money it invests in schools is being used.'

Money and output are intrinsically linked but the *measures* of output would seem the debatable question. Is the output justifying the amount of expenditure invested in curriculum and examinations work? Of course this does depend on what can actually be measured in education. The revised National Curriculum which came on line for September, 1995 resulted in all previous curriculum documents being superseded. Is the cost of devising annual tests with questionable monitoring potential worth the investment of money to produce the tests and the time involved in administering them? The total operating costs for the School Curriculum and Assessment Authority(SCAA) for 1994-95 was £30.4 million (SCAA: 1995:5-6) of which £20.8 million or 68.4% accounted for curriculum, assessment and examinations work. The costs of tests for Key Stage 1 works out at £2.6 million or 12.5% of the budget for curriculum, assessment and

examinations. This represents quite an investment of public revenue. How can it be justified?

Allocation of funding to schools is not always to the level needed and does vary from county to county in England. Vast sums of money in monitoring must be seen to be justified to the profession. There is considerable doubt as to such justification. It is also somewhat daunting to teachers who are responsible for fulfilling policy which is imposed without sufficient debate. It seems pointless to test children on tests which provide little information to the overall benefit of a child's learning. It is questionable when national tests need modifications year after year to make them manageable educationally, financially and politically.

The concept of quality in education is certainly under question. One has only to look at the term 'Quality Assurance' and the structures which schools and institutions are now putting in place to monitor and demonstrate the quality of output to understand the emphasis placed on ascertaining value for money.

As I write this paper, a Report from the Chief HMI indicates that half the primary schools and two-fifths of secondary schools in England are failing to teach children to a satisfactory standard. Three main areas of concern were identified in relation to Literacy, Numeracy and Teaching Quality. The OFSTED report drew conclusions from the reports of more than four thousand inspections. Attention was drawn to a 'worryingly persistent' slowing of pupils' progress at ages seven and eight. 'This dip in performance is strongly associated with a fall in the quality of teaching.'. Schools need to be more exacting in tracking pupils' progress and monitoring teaching. The main points which emerged from the report were:

* One in three lessons for 8 to 11 year olds were unsatisfactory;
* One in five lessons for 11 to 14 year olds were unsatisfactory;
* Two out of five lessons are good or very good;
* One in ten primary schools need to raise reading standards for 11 year olds;
* A good start at Key Stage 1 is lost at key Stage 2;
* Primary schools should try harder to track pupils' progress;
* Higher standards should be set in literacy with more time spent on phonics;
* Two in five schools should make better use of resources;
* One in 12 secondary schools and one in 17 primaries are seriously short of books and equipment;
* Estimated 48,000 teachers 'excellent'; 15,000 'poor'.

Immediate action was necessary for literacy to tackle mediocre and poor standards. This was most apparent for children between 8 and 11. Three-quarters of seven-year olds reached or exceeded level 2 which is seen as an appropriate level for stretching children's reading, writing and spelling at that age. One-quarter of children not achieving level 2, for whatever reason was felt to be still too high a number. Data and interpretations from newspaper coverage can be hard-hitting as the following example illustrates with figures moving from a quarter to a third of children:

> 'The failure of primary schools to teach one child in three to read properly by the age of seven is the education system's greatest flaw. It is entirely inexcusable and the consequences are devastating, as much for the individual as the community.' (Clare 1995:5)

Intertwined with poor results is the issue of the most appropriate approach to teach reading. The teaching of phonics it was claimed was ' too rarely an established part of a well-structured program for all pupils.'

> 'Approaches to the teaching of reading generally include some phonic work. The place and purpose of teaching phonics, however, rarely feature strongly in school reading policies. Consequently, the teaching of phonic skills is not as thorough as it should be.'

Literacy standards are frequently brought into question and it is not surprising that different perceptions of the state of literacy cause confusion. The above picture from OFSTED data contrasts with a 1995 NFER (National Foundation for Educational Research) survey published two months previously indicating that reading standards for eight year old children had improved significantly since a dip in 1991. An inference drawn by the NFER indicated that' the disruption primary schools were put through at the end of the Eighties and into 1991 would have contributed to the decline that

was measured then'. (Blackburne 1995:2) Crucial to the debate on literacy are factors such as reading approach, classroom organization and management. If a literature approach is used to teach reading to very young children then the one-to-one approach of hearing reading can take place across the school day when children can be working on other curriculum areas. Figures from the OFSTED (Office of Standards in Education) report indicate that 'about seven-tenths of good lessons involve a balance of whole class teaching and group work. Poor lessons, conversely, often lack this pragmatic balance'. Research and surveys in recent years seem to produce conflicting evidence which serves only to add confusion as to whether or not standards are declining or increasing. A link between OFSTED inspection data and the results of Key Stage 1 reading has not as yet been explored to see if what the inspection teams are finding correlate with results. More research seems needed to examine the range of reading approaches in relation to the curriculum so that approaches, methods and results can be seen from a more informed baseline. The enormous wealth of data produced by the OFSTED inspections will be a potentially rich source of information for research but this will need to be tempered with other research so that a more accurate idea of literacy teaching and learning can emerge. The NFER had funded two yearly surveys into reading standards and would like Government funding to continue what is seen as a more independent approach to researching and monitoring reading in schools.

Current concern for literacy standards has seen the injection of £5 million available for Local Education Authorities to bid for grants in a 1996-97 Government project which aims to raise standards of literacy and numeracy in schools. The project aims to establish a network of 20 centers in co-operation with OFSTED to work with groups of primary schools. It is envisaged that the project will examine OFSTED evidence about effective teaching and learning in reading with in-service help for teachers being made available.

However the variables between schools are considerable. Primary schools are usually small institutions reflecting the needs and composition of the area they serve- whether inner city, suburban, town or rural. The success or failure of a few children can skew results on tests when comparisons are made and if league tables become established. This is tricky when resource distribution may depend on results. It does not mean that expectations in disadvantaged schools should not be high. Leaving aside environment issues related to each school, the financial base at the school level would seem to vary considerably. Howlett(1992: 20 )comments on a Coopers & Lybrand survey for the National Union of Teachers in estimating how much it would cost an average primary school to implement the National Curriculum.

> 'The widely differing financial positions of individual schools made it difficult to generalize about the need for spending on books, consumables and equipment'.

Somehow the practical 'extras' to implement a curriculum change such as the National Curriculum are never included. Extra direct costs for a primary school, for example, were estimated by Coopers & Lybrand to approximate 6% on the average school budget. But as budgets were generally tightened, the extra money needed did not eventuate.... reality of 'more for less' meant absorbing new initiatives. This is perhaps one example illustrating the 'invisible' holes which characterize change.

The quest for accountability be it through more curriculum control and assessment procedures particularly focuses on children at identified phases in education. In England, Primary Education is being examined carefully; nursery education and is settling on the political agenda with increased public awareness and demand for greater provision. A recent report by the Prime Minister to introduce targets for four year old children as a basis for measuring children's progress illustrated current thinking in relation to monitoring the progress of the very young.

> 'When the (pre-school targets) are in place, we will have a simple baseline testing at the start of primary school, which will put in place the foundation stone I want to see for testing at all key stages in school life'. (Blackburne 1995:6)

It looks as though children's progress will be measured earlier and earlier. While the early years are important for children's cognitive and social development, the need to measure knowledge and skills by the age of 5 and, at so early an age is controversial, particularly at the national level. Provision for pre-school education is quite patchy in the UK and variable in quality. A situation could arise with certain children taking tests while others had not even started preschool education. Such baseline testing undertaken by teachers would provide the useful diagnostic information to help

children's learning. National monitoring of pre-school targets would only produce the rather global assessments that are inherent in the Key Stage 1 SATs for seven year old children. As the type of tests and results for Key Stage 1 are questionable, considerably more research into testing pre-school children is needed. Also, the participation of those working in this area is essential so that the tests reflect informed ideas for teaching and assessing children at so young an age.

Once again, the quality of learning and value for money are significant influences. At a recent conference held in the north of England, Gillian Shepherd, the Education and Employment Secretary indicated the importance of value for money in the Government's scheme to give parents of four -year-olds vouchers worth £1,100. The government intended to invest £750 million in the voucher scheme. 'It must be seen to result in improved standards. Value for money is necessary. It is our duty to assess improvement.' (Blackburne 1996:2)

The impact of the National Curriculum and the related tests is filtering down the educational ladder. A consequence could well be a pre-school National Curriculum. The pre-school goals identified were narrow in one sense but broad in another.
For example, the goals for Language and Literacy are worded as follows:

> 'In small and large groups, children listen attentively and talk about their experiences. They use a growing vocabulary with increasing fluency to express thoughts and convey meaning to the listener. They listen and respond to stories, songs, nursery rhymes and poems. They make up their own stories and take part in role play.
>
> Children enjoy books and handle them carefully, understanding how they are organized. They begin to associate sounds with patterns in rhymes, with syllables, and with words and letters. They recognize their own names and some familiar words. They recognize letters of the alphabet by shape and sound. In their writing they use pictures, symbols, familiar words and letters, to communicate meaning, showing awareness of some of the different purposes of writing. They write their names with appropriate use of upper and lower case letters.'

Use of words such as 'attentively', 'growing', 'some', 'awareness' are assumed to have a consensus but such interpretation depends greatly on context and the parameters of understanding. This is then quite a formidable task when quantifying young children's learning from such a broad base. Conversely, the itemizing of national objectives or learning outcomes to test literacy would also be criticized as being too prescribed. In many ways this is the dilemma facing the revised National Curriculum and the testing program. To what extent they complement one another is highly debatable.

## QUESTIONS OF PROCEDURE

For tests to have credibility, points of procedure must be exacting. '
'For each subject a test and task specification was drawn up by SCAA and agreed with the Department for Education and Employment. This provided the framework to which test writers worked. The test questions and task activities were drafted by experienced writers together with practicing classroom teachers.' (SCAA 1995:11) While the approach would seem a consultative one there are many issues concerning time, format, content to represent age, interest, linguistic aspects of language and expediency to mention just a few. Trials were undertaken in schools with modifications made from children 's reactions and teachers' thoughts. Formal trials, for the 1995 SATs were done in June and October 1994. The sample size of over 6,000 children represented a range of schools across England.

SATs arrive in schools quite some weeks before the suggested dates of administration. The tests rely heavily on an element of trust. For key Stage 1, the timing of the tests seems to be remarkably lengthy.

'The assessment period for the tasks has been extended so that the tasks may be completed at any time between the beginning of January 1996 and four weeks before the end of the Summer term. This brings the start date of tasks at Key Stage 1 into line with Key Stages 2 and 3.'(SCAA 1955:3)

Recognized essentials of validity and reliability found in standardized tests seem remarkably absent in the model of SATs. While sample sizes and geographical location would appear representative, elements of doubt seem to arise from the type of test that has emerged especially points of the procedure which are difficult to control. In calling the tests, tasks, in no way softens the overall intention to monitor progress. For educational and presumably policy decisions to be made, then a high degree of consistency must prevail in test construction and implementation.

An examination of the percentage of children who gave correct answers to the 1995 Level 3 Reading Comprehension Test 'Holes and Tunnels' shows quite a disparity in relation to progression of individual question difficulty. For the story on 'Vicky's Wobbly Tooth' the first question gave 90% response. Question 2 shows a drop to 59%; Question 3 revealing a dramatic drop to 28% then an increase for two questions with a final question at 51%.

Questions that required factual answers were handled better than those that asked for inference and deduction. Questions that repeated the wording of the text were easier that those which required children to change the wording; those questions requiring identification of main facts or the gist of a passage were easier than those demanding close reading.

The 1995 evaluation indicated that the vast majority of teachers felt comfortable with the tasks and found them 'manageable, suitable for the age range, and appropriate in terms of curriculum approach.' (SCAA 1995: 6). A quarter of teachers in the sample rated the validity of the Level 1 Reading Task in assessing reading as 'less than satisfactory'. (SCAA 1995:6) It is acknowledged in the evaluation that the assessment of children at early stages of reading development presents a 'special challenge'.

The Level 3 Reading Comprehension Test was rated satisfactory by 74% of teachers in respect of manageability and curriculum approach. There were reservations expressed by a small number of teachers as to the 'suitability and validity of the test as a means of assessing Reading at Key Stage 1' (SCAA 1995:7) The test it was felt covered only a part of the Attainment Target for Reading at Level 3.

## ASSESSING CHANGES

The SATs have been modified each year to include changes through recommendations made either from consultation or teacher pressure. This is inevitable in an approach to assessment which may be termed 'evolutionary'. While the benefits may pay off in the long term, the yearly amendments make points of comparison in determining standards rather difficult and perhaps impossible. In 1995, the changes for English at Key Stage 1 were to initiate a more consistent format for reading comprehension; to do away with a separate spelling test and to provide a challenge for more able students by allowing Level 4 pupils to be assessed with Key Stage 2 material. In 1995, able children 'could demonstrate their performance by taking tests developed for pupils at the end of key stage 2 '. (SCAA 1994:8) Other changes allowed grants to LEAs to fund supply cover to help teachers who were testing children in the classroom; more guidance in the handbook to provide more flexibility to teachers in setting the tasks; changes in the layout of the printed materials.
SCAA would view the tests as reflecting a broad range of curriculum contexts and are therefore more reliable than standardized tests. Reliability and validity are essential to the worth of a good standardized test but, in respect to SATs these essentials are certainly questionable. (Hofkins 1996:5) reporting on the thoughts of SCAA states that:

'One possibility is to convert pupils' raw marks, used to determine what national curriculum level they are at, into standardized scores'.

Are the SATs then nearing the crossroads of a merger between criterion-referenced tests and standardized tests? Is the standardized score more manageable for statistical purposes in giving a quantitative indicator of a child's performance.

Also, the SATs are not taking a child's age into account. A standardized score, it might be argued, would not disadvantage a younger child. Plans will now go ahead to produce standardized scores from the 1996 tests in reading comprehension.

As a measure of children's performance the tests are limited by the skills and knowledge that can be realistically tested. This, however applies to most tests. The SATs in its first version tested more than the leaner and refined later versions. Each refinement has a resulting limitation in what is measured to show a child's performance. Modifications to the tests and the measurement implied in monitoring could well see the introduction of a numerical standard.

## NATIONAL LEAGUES

What worries the teaching profession is ranking schools on merit from the results obtained across particular subjects on tests which have weaknesses and the consequent interpretation which the public makes of the results. Having established a secondary school league, movements are now underway to introduce one at primary level. The first such publication is likely to be on the tests for eleven year old children to be conducted in the summer of 1996. The first major run of these tests was in 1995 and teachers were concerned by the validity of the tests and the type of information on which judgments are made. Will a league descend from upper primary to lower primary as more pressure is placed on performance? And then to nursery? Where is the bottom line?

## PROBLEMS

Establishing baselines from such an approach takes time. Now that Key Stage 1 tests have been implemented for five years, data is developing. The changes to the tests since their introduction in 1991 do make exact comparisons impossible. Those tests administered in 1995 may provide the first baseline information because not every school administered the SATs in 1993 and 1994 due to a boycott.

At Key Stage 1 the results show an improvement over the years; eighty percent of children reached level 2 with an increase in those reaching level 3. How reliable were the original baselines...or how were the original baselines devised? Tooley (1996:20) recounts how the Government's National Curriculum working groups 'got it wrong when they guessed the levels of attainment for certain age groups' and recalls the work of TGAT(the Task Group on Assessment and Testing) set up to advise on a testing program. It suggested a 10-level scale to chart progress for children of different ages. Agreement by the committee on levels of attainment from average children taxed the committee for six months without any progress before a format evolved within the year's deadline given by the then Secretary of State for Education. Tooley is not surprised that the expected levels are inexact 'given the paucity of research available, the tight deadlines and the political pressures they worked under'.

Gremlins in new tests take a little while to disappear. More so when modifications produce new gremlins! Thousands of lower-ability children and those with special needs who undertook the Key Stage 2 tests last year were double counted in the results published by the Government in January, 1996. This will be amended for 1997.

The tests for seven -year-olds claim researchers from Manchester University are inaccurate and misleading. 'Children who scored highly in the tests were barely able to pass internationally recognized reading levels. Although school tests suggested seven-year-olds improved every year, more rigorous examinations indicated falling standards.' (Scott-Clark 1996:7)

To what extent the tests actually help teachers would seem quite debatable in diagnosing children's needs. The tests seem more designed to national monitoring of a small part comprising literacy and serve to highlight school and teacher weaknesses in relation to the knowledge and skills tested. Also, teachers moved relatively quickly from having complete autonomy over the curriculum with support of Local Education Authorities to a National Curriculum and national test

program. A political implication was to diffuse the power of Local Education Authorities. While power did exist at local level it was not as highly centralized as the developing power of national government.

## INTERNATIONAL UNDERSTANDING

It almost seems that there is a national isolationism with respect to appropriate approaches to testing literacy. France seems to have developed more diagnostic tests which are designed to identify pupils' needs rather than teachers' weaknesses. Current approaches to testing require children at 8, 11 and 15 to be tested at the beginning of the school year have been in operation since 1989. From their position in the school year the tests are in place to enable teachers to diagnose rather than evaluate in a summative capacity. The results are not used to compare schools or to evaluate teacher performance.

The developing model of samples of work in a portfolio format will provide teachers with an example of qualitative assessment and is an excellent idea but quite experimental as the model moves into a quantitative form to monitor progress nationally. Apart from the common procedures necessary, a highly sophisticated system of internal and external moderation would need to be in place so that the portfolio approach serves the different audiences for accountability.

## CONCLUSIONS

One of the significant lessons emanating from the experience in England and Wales is the importance of having *teachers as part of the process.* This idea is not new and would seem almost a pre requisite. There is much school-based curriculum development work dating from the 1970's to suggest that participation and commitment are essential elements. Somehow this experience has not been tapped by those responsible for implementing government policy. Is it that government policy does not want such collaboration? Curriculum must have teacher involvement if it is to work; consultation is no substitute. Similarly, the more complex area of assessment must also reflect involvement. This point emerged from the 1996 British Council/SCAA seminar by participating countries was that testing could only be introduced successfully if governments had the support of teachers.

The revised National Curriculum with 'level descriptions' for teaching and learning is a step forward in putting the *child* at the *center* of the assessment process and not the task. During the Dearing review much dissatisfaction was directed towards the 10-level framework for SATs at all key stages. Teacher associations, teachers and all those who participated in the consultations were evenly divided as to the retention of the levels. As no viable alternative came to the fore, and a framework had to be in place by a particular time, expediency won.

Government recognition that primary *teachers need more help with the complexities in assessing children's progress* is indeed another welcome step. Rather than simply ticking boxes which represented attainments, the revised approach requires teachers to follow a 'best fit' model. This relies more on teacher judgment and is more subjective in ascertaining a child's achievement. Questions of objectivity in the absence of a rigorous system of moderation may cause concern and question reliability of school results.

Once again, the new system was introduced as a development indicating the inherent difficulties of an evolutionary model. Training needed to be in place for the launch of so important a change. This was not to be.

It might be seen that *assessment has come full circle* during the past eight years.....moving from teacher assessment to measuring by statements of attainment and back to teacher judgment. Such a simplistic view does not take into consideration the need for an organized approach to involve school, LEA and national government to establish a qualitative and quantitative yardstick.

Across the years there has been an increasing *centralization in curriculum and assessment for political and economic reasons.* Traditional curriculum autonomy assigned to the school has disappeared despite the devolving of budgets to schools.

Curriculum autonomy may not have been lost had primary teachers had *a professionally agreed curriculum and assessment procedure in place* for the past decade along with a professional voice to champion the quality of teaching and learning in primary schools.

How much national assessment is necessary and *where does one draw the line?* The original points of ages 7, 11 and 14 for tests are now established. Will pressure lower the age to five? Will the idea for 'optional' tests at age nine become a reality in the light of poor results obtained at Key Stage 2 and the suggestion by SCAA that more formal assessment half-way through Key Stage 2 would enable teachers to more effectively track children's progress. (Hofkins 1996:13)

The experience in England is showing how an evolutionary approach to assessment with identified levels can seen by Government to represent national standards. The journey since 1987 has been long and expensive with many lessons to learn.

An international forum such as this does enable views to be exchanged. Individual associations from different countries concerned with literacy could contribute their expertise in developing approaches which serve the needs of children, teachers, administrators and governments. While contexts of countries and education systems do vary, there perhaps are many principles and ideas which are common. To what extent the voice of national literacy organizations are listened to no doubt varies. There is a wealth of experience and knowledge internationally on assessing reading. This ought to be harnessed to refine national assessment to alleviate the expensive trial and error initiatives and numerous surveys which over the years have done little to implement a co-ordinated approach that reflects true accountability.

---

**BIBLIOGRAPHY**


Australian Literacy Federation *Common Ground,* The Newsletter of the Australian Literacy Federation, No.3 November,(1995), p.1

BLACKBURNE, L. 'Testing vow alarms early-years world' in *The Times Educational Supplement,* Sept 22,(1995), p.6.

BLACKBURNE, L. 'Survey points to improved reading' in *The Times Educational Supplement,* Dec 15,1995, p.2.

BLACKBURNE, L. 'National tests loom for five-year-olds' in *The Times Educational Supplement,* Jan 12,(1996), p.2.

CLARE, J. 'Almost all Sue Lloyd's pupils can read properly but a third of children in England cannot. Why?' in *The Daily Telegraph,* Nov 28, (1995), p.5.

DES *Task Group on Assessment and Testing(TGAT): A Report,* London: Department of Education and Science, 1987.

HAIGH, G. 'Testing this year's models' in *The Times Educational Supplement,* Feb 10,(1995), p.5.

HOFKINS, D. 'Tests could be made more useful' in *The Times Educational Supplement,* Feb 2,(1996) p.5; p.13.

HOWLETT. P. 'An Annual Expense in *The Times Educational Supplement,* Jan 17, (1992) p.20.

SEITZ, R. 'Letters from over here' Radio Broadcasts, BBC Radio 4, Jan-Feb, 1996.

SCAA *Assessing 7 and 11 Year Olds in 1995,* London: SCAA ,1996

SCAA *Annual Report: April 1994- March 1995* London; SCAA, 1995 pp. 5-6.

SCAA *Report on the 1995 Key Stage 1 Tests and Tasks n English and Mathematics,* London: SCAA, 1995.

SCAA *Key Stage 1 Assessment Arrangements 1996* London: SCAA/ DfEE, 1995

SCAA *Desirable Outcomes,* London: SCAA/DfEE, 1996.

SCOTT-CLARK, C. & HYMAS, C. 'Former pupils sue over their bad education' *The Sunday Times,* Feb 25,(1996) p.7.

TOOLEY, J. 'Lies, damn lies and government statistics' *The Times Educational Supplement,* Feb 2,(1996) p.2.

WOODHEAD, C.(1996) 'When outsiders know best' *The Guardian,* Education section, Feb 6, (1996) p.2.

---

**INTERNATIONAL LITERACY INSTITUTE, University of Pennsylvania, 3910 Chestnut Street, Philadelphia, PA 19104-3111 USA; Telephone 215-898-2100; Fax 215-898-9804;**
**E-mail: ili@literacy.upenn.edu**
**URL: http://www.literacyonline.org**